

AD-A044 573

CARNEGIE-MELLON UNIV PITTSBURGH PA DEPT OF STATISTICS

F/G 7/4

A NEW STOCHASTIC ANALYSIS OF CHEMICAL KINETICS.(U)

JUN 77 J P LEHOCZKY

AFOSR-74-2642

UNCLASSIFIED

TR-129

AFOSR-TR-77-1212

NL

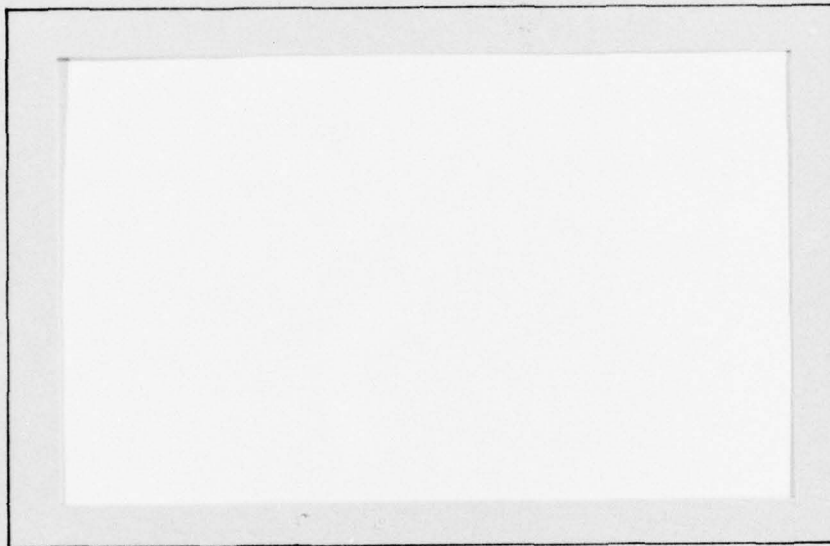
| OF |

AD
A044573



AD A 044573

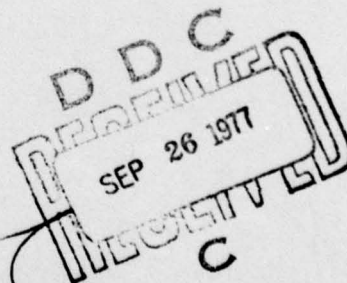
AFOSR-TR- 77- 1212



DEPARTMENT
OF
STATISTICS

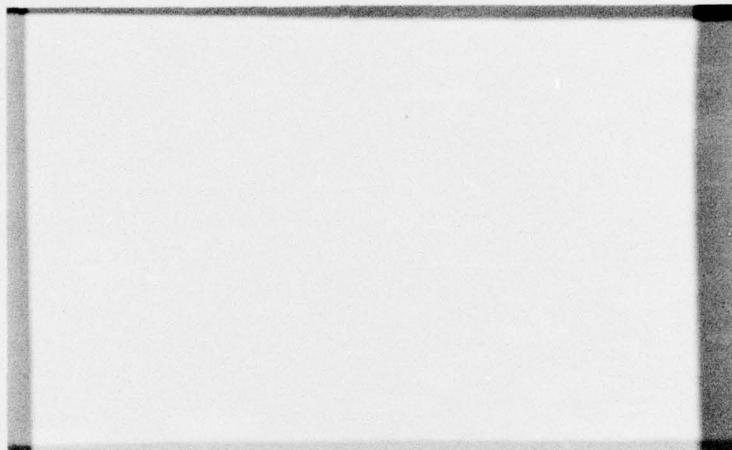
AD No. _____
DDC FILE COPY

Approved for public release;
distribution unlimited.



Carnegie-Mellon University

PITTSBURGH, PENNSYLVANIA 15213



AIR FORCE OFFICE OF SCIENTIFIC RESEARCH (AFSC)
NOTICE OF TRANSMITTAL TO DDC
This technical report has been reviewed and is
approved for public release IAW AFR 190-12 (7b).
Distribution is unlimited.
A. D. BLOSE
Technical Information Officer

A NEW STOCHASTIC ANALYSIS OF
CHEMICAL KINETICS

by

John P. Lehoczky*
Carnegie-Mellon University

June, 1977

Technical Report No. 129

Department of Statistics
Carnegie-Mellon University
Pittsburgh, PA 15213



* Research supported in part by Grant AFOSR 74-2642b
from the Air Force Office of Scientific Research.

A NEW STOCHASTIC ANALYSIS OF CHEMICAL KINETICS

by

John P. Lehoczky*
Department of Statistics
Carnegie-Mellon University
Pittsburgh, PA 15213

ABSTRACT

A new approach to the formulation and analysis of stochastic models of chemical reactions is presented. Unimolecular, bimolecular, and enzyme kinetic reactions are considered in the irreversible and reversible cases. The methodology is based on diffusion approximations and represents the time evolution of the reaction as the sum of a deterministic function and an Ornstein-Uhlenbeck process. As a result the marginal distributions are approximately Gaussian with relatively simple mean and covariance parameters, and the dynamic behavior is completely characterized. The stochastic approach which uses stochastic differential equations is a natural generalization of the deterministic approach which uses ordinary differential equations.

Y for	White Section	<input checked="" type="checkbox"/>
J	Buff Section	<input type="checkbox"/>
DDC	UNANNOUNCED JUSTIFICATION	
BY	DATE	

A

* Research supported in part by Grant AFOSR 74-2642b from the Air Force Office of Scientific Research.

1. Introduction

In this paper a new approach to the formulation and analysis of stochastic models of chemical reactions is presented. Unimolecular, bimolecular, and enzyme reactions are considered in both the irreversible and reversible cases. The reaction types chosen for analysis are of a simple sort; however, this is for purposes of exposition only, and the methodology can be easily extended to very complicated reactions. The results developed in this paper should be compared for their simplicity with those obtained by other methods. The survey of McQuarrie^[1], ^[2] provides an excellent summary of such results.

Early development of the stochastic approach to chemical kinetics is due to Singer^[3] and Bharucha-Reid^[4]. This approach to chemical kinetics typically treats the concentrations of various molecule types as a continuous time Markov chain with specified transition rates. The transition rates are used to construct the Kolmogorov forward equations which characterize the probabilistic behavior of the reaction. The exact solution of these partial differential difference equations is intractable except in the simplest case, so most authors are content to find the first two moments of the process. Even this often leads to differential equations whose solutions entail special mathematical functions. In such cases it is difficult to gain an intuitive understanding of the process. Furthermore, such an analysis is very limited in that it provides information only about moments and not the exact distribution. Moreover, it provides no insight into the dynamic behavior or n-dimensional distributions over time. Finally, as the reaction becomes more complicated

the analysis becomes more intractable and the insights less available.

The approach presented in this paper is approximate rather than exact, but overcomes all of the difficulties which arise with a stochastic analysis mentioned before. Furthermore it provides a clear connection between the usual deterministic analysis and a stochastic analysis. The method is called a diffusion approximation, and examples of its use are given in the work of Gaver and Lehoczky^{[5], [6], [7]} and McNeil and Schach^[8]. The approach is asymptotic in that it assumes the number of molecules involved in the reaction is large, although in the work of Gaver and Lehoczky previously cited numerical studies indicate excellent accuracy in small systems as well. The fundamental idea is based on the replacement of the discrete state space by a continuous state space and the Markov chain by a Markov process (in these cases a diffusion process). The forward equations (in the Markov process case known as the Fokker-Planck equations) are then formulated as Ito type stochastic differential equations rather than their equivalent version as partial differential equations. This formulation permits the use of the Ito calculus for the manipulation of stochastic differential equations. The reader may consult Arnold^[9] or Gikhman and Skorokhod^[10] for a treatment of this theory. The stochastic process is, in the asymptotic case, transformed into the sum of a deterministic process (invariably the same as the process arising from a deterministic analysis of the system) and a stochastic process (invariably an Ornstein-Uhlenbeck process).

This characterization of the system provides a full description of the marginal distribution (Gaussian with specified mean and covariance functions) over time and, more importantly, of the n-dimensional distributions over time (Gaussian with specified mean vector and covariance structure). The characterization of the n-dimensional distribution functions is important for the development of a proper statistical analysis of data gathered from such a process say to estimate the rate constants of the reaction. Since the stochastic process is added to the deterministic process, the deterministic process may be thought of as a first order approximation.

It is important to note that the methodology given in this paper can easily be applied to very complicated reactions including those with many compounds and steps. The results are relatively simple and easily understood. Once the nature of the Ornstein-Uhlenbeck process is understood, its appearance in the description of many difference reactions allows for greater insight into chemical processes.

2. Unimolecular Reactions

In this section, two simple unimolecular reactions are considered. The two selected are taken from the review by McQuarrie^[1] in order to provide a basis of comparison for the diffusion approximation results with the exact results.

A. Reaction $A \xrightarrow{k} B$ [11],[12],[13],[14]

If $A(t)$ represents the number of A molecules present at time t , then $\{A(t), t \leq 0\}$ is typically treated as a continuous time Markov chain with transitions given by

<u>Transition</u>	<u>Rate</u>	
$a \rightarrow a - 1$	$kadt + o(dt)$	
$a \rightarrow a$	$1 - kadt + o(dt)$	2.A.1
all others	$o(dt)$	

where $o(h)/h \rightarrow 0$ as $h \rightarrow 0$.

For large values of $A(0)$, the time between transitions is very small (having a mean value of $k/A(t)$). Consequently, over a short period of time many transitions will occur. It is therefore reasonable to assume $A(t+dt) - A(t) = dA(t)$ has approximately a Gaussian distribution with mean $E(dA(t)) = -kadt + o(dt)$ and variance $\text{Var}(dA(t)) = kadt + o(dt)$. One may therefore represent $A(t)$ as the solution of the Ito stochastic differential equation (s.d.e. in the sequel)

$$dA(t) = -kA(t)dt + (kA(t))^{1/2} dW(t) \quad 2.A.2$$

where $\{W(t), t \geq 0\}$ is a standard Wiener process and $dW(t) = W(t+dt) - W(t)$.

The state space for the Markov chain was $\{0, 1, \dots, A(0)\}$ originally and is now $[0, A(0)]$. The infinitesimal mean and variance in (2.A.2), $-kA(t)dt$ and $kA(t)dt$, were chosen to match that of the original Markov chain keeping terms of order dt only. The approximation of the discrete process by a continuous process is made rigorous in the work of Kurtz^[15] and Barbour^[16]. The increments are modelled as having a Gaussian distribution in view of the central limit theorem. In the spirit of the central limit theorem, it is natural to introduce another stochastic process $\{X_N(t), t \geq 0\}$ where $X_N(t) = (A(t) - Na(t))/N^{1/2}$ with $N = A(0)$ and $a(t)$ is an arbitrary deterministic function. The X_N process will also be a diffusion and will satisfy a s.d.e. which can be determined using Ito's Lemma [9,p.70] or [10,p.27]. Specifically,

$$\begin{aligned} dX_N(t) = & -N^{1/2}(a'(t) + ka(t))dt - kX_N(t)dt \\ & + (ka(t) + O(N^{-1/2}))^{1/2} dW(t). \end{aligned} \quad 2.A.3$$

As $A(0) = N \rightarrow \infty$, the $N^{1/2}$ term in (2.A.3) will explode unless its coefficient is exactly 0. The function $a(t)$ must, therefore, be chosen to satisfy the equation

$$a'(t) = -ka(t), \quad a(0) = 1. \quad 2.A.4$$

Given that $a(t)$ satisfies (2.A.4) and therefore equals $a(t) = \exp(kt)$, one may let $N \rightarrow \infty$ in (2.A.3) and conclude from Barbour^[16] Theorem K that $\{X_N(t), t \geq 0\}$ converges weakly to $\{X(t), t \geq 0\}$ where X satisfies the s.d.e.

$$\begin{aligned} dX(t) &= -kX(t)dt + (k \exp(-kt))^{1/2} dW(t), \\ X(0) &= 0. \end{aligned} \quad (2.A.5)$$

The diffusion approximation consists of writing

$$A(t) \equiv Na(t) + N^{1/2}X_N(t) \approx Na(t) + N^{1/2}X(t). \quad (2.A.6)$$

The stochastic process X is a nonstationary Ornstein-Uhlenbeck process with mean 0. Many facts about X can be deduced from this characterization including the joint distribution functions, the moments, and the spectrum [9, Chapter 8]. The X process will have mean 0 and variance Σ_t given to be the solution of

$$\dot{\Sigma}_t = -2k\Sigma_t + k \exp(-kt), \quad \Sigma_0 = 0. \quad (2.A.7)$$

Equation (2.A.7) is readily solved to give

$$\Sigma_t = \exp(-kt)(1 - \exp(-kt)). \quad (2.A.8)$$

The X process is Gaussian and $X(t)$ has a Gaussian distribution with mean $\exp(-kt)$ and variance $\exp(-kt)(1 - \exp(-kt))$, thus $A(t)$ will have approximately a Gaussian distribution with mean $A(0)\exp(-kt)$ and variance $A(0)\exp(-kt)(1 - \exp(-kt))$. This approximation is clearly in accordance with the exact results for this process, namely a Binomial $(A(0), \exp(-kt))$ distribution, as the central limit theorem guarantees the goodness of our approximation even for small values of $A(0)$. The diffusion approximation (2.A.6) provides a complete description of the dynamic behavior of the reaction rather than merely describing the static or marginal behavior. The approximation is motivated

by the large $A(o)$ case, but the reader is referred to the work of Gaver and Lehoczky [5], [6], [7] for more information concerning the accuracy of the method. Excellent accuracy has been reported for values of $A(o)$ as small as 20.

B. Reversible Unimolecular Reaction $A \xrightleftharpoons[k_2]{k_1} B$, [1], [13], [14], [17], [18], [19].

The irreversible reaction of section 2A is now generalized to the reversible case. The forward and backward rate constants are k_1 and k_2 respectively. The transitions and transition rates appropriate for such reactions are given by

<u>Transition</u>	<u>Rate</u>
$a \rightarrow a+1$	$k_2(B(o) + A(o) - a)dt + o(dt)$
$a \rightarrow a-1$	$k_1 a dt + o(dt)$
$a \rightarrow a$	$1 - (k_1 a + k_2(B(o) + A(o) - a))dt + o(dt)$
all others	$o(dt)$

2.B.1

The A process can again be approximated by the s.d.e.

$$dA(t) = k_2(B(o) + A(o) - A(t))dt - k_1 A(t)dt \quad 2.B.2$$

$$+ (k_2(B(o) + A(o) - A(t)))^{1/2} dW_1(t) - (k_1 A(t))^{1/2} dW_2(t)$$

where $W_1(t)$ and $W_2(t)$ are independent standard Wiener processes. An X_N process is again introduced where $X_N(t) = (A(t) - Na(t))/N^{1/2}$ and $N = A(o) + B(o)$. Let $a_o = A(o)/N$ and $b_o = B(o)/N$. Using Ito's Lemma one may derive the s.d.e. for the X_N process to be

$$dX_N(t) = -N^{1/2}(a'(t) - k_2(1 - a(t)) + k_1 a(t)) \\ - (k_1 + k_2)X_N(t)dt + (k_2(1 - a(t)) + k_1 a(t) + O(N^{-1/2}))^{1/2} dW(t). \quad 2.B.3$$

Letting $N \rightarrow \infty$, the coefficient of the $N^{1/2}$ term must be 0, so $a(t)$ must satisfy

$$a'(t) = -k_1 a(t) + k_2(1 - a(t)), \quad a(0) = a_0 \quad 2.B.4$$

or

$$a(t) = (a_0 - \frac{k_2}{k_1 + k_2}) \exp(-(k_1 + k_2)t) + \frac{k_2}{k_1 + k_2}.$$

If $a(t)$ is given as above, then $\{X_N(t), t \geq 0\}$ converges weakly to $\{X(t), t \geq 0\}$, a diffusion process which satisfies

$$dX(t) = -(k_1 + k_2)X(t)dt + (k_2(1 - a(t)) + k_1 a(t))^{1/2} dW(t) \quad 2.B.5$$

$$X(0) = 0.$$

The X process is a nonstationary Ornstein-Uhlenbeck process, and the diffusion approximation methodology treats $A(t)$ as being given by $(A(0) + B(0))a(t) + (A(0) + B(0))X(t)$. One may easily compute the variance of the approximation process, Σ_t , to be the solution of

$$\dot{\Sigma}_t = -2(k_1 + k_2)\Sigma_t + k_2(1 - a(t)) + k_1 a(t), \quad \Sigma_0 = 0. \quad 2.B.6$$

The solution is given by

$$\Sigma_t = \exp(-Kt)(1 - \exp(-Kt)) \left\{ \frac{k_1 k_2}{K^2} (\exp(Kt) + 1) \right. \\ \left. + (k_1 - k_2)(a_0 - k_2/K) \right\} \quad 2.B.7$$

where $K = k_1 + k_2$.

One may let $t \rightarrow \infty$ to deduce the equilibrium distribution of $A(t)$. In this case $A(t)$ will have a steady state Gaussian distribution with mean $k_2(A(o) + B(o))/K$ and variance $k_1 k_2 (A(o) + B(o))/k^2$. These results extend those previously obtained for this reversible reaction. Of particular importance is the characterization of the dynamic behavior in terms of an Ornstein-Uhlenbeck process. The n -dimensional distribution functions over time are thus given by an n -dimensional Gaussian distribution.

Many other unimolecular reactions can be treated in this manner including the parallel first order and triangular reactions described by McQuarrie^[1]. Instead we next treat the more complicated nonlinear situations.

3. Bimolecular Reactions

The bimolecular reactions to be considered are again summarized in the survey of McQuarrie^[1]. The methodology of diffusion approximations is especially useful, because it provides simple characterizations of the evolution of these processes in terms of the sum of a deterministic function and an Ornstein-Uhlenbeck process. This is in sharp contrast to the results obtained from an exact analysis. Such results, when they can be obtained at all, invariably involve complicated combinations of special mathematical functions. Consequently, it has been impossible to obtain any intuitive understanding of the stochastic nature of the chemical reaction.

A. The Irreversible Reaction $2A \xrightarrow{k} B$

A number of asymptotically equivalent stochastic models are possible. If $A(t)$ represents the concentration of A molecules at time t , then $\{A(t), t \geq 0\}$ is taken to be a Markov chain with transitions and transition rates given by

<u>Transition</u>	<u>Rate</u>	
$a \rightarrow a - 2$	$\frac{ka^2}{2(A(o) + B(o))}$	$dt + o(dt)$
$a \rightarrow a$	$1 - \frac{ka^2}{2(A(o) + B(o))}$	$dt + o(dt)$
all others	$o(dt)$	

3.A.1

Using (3.A.1) and defining $dA(t) = A(t+dt) - A(t)$, it is easy to show $E(dA(t)) = -kA^2(t)dt/(A(o) + B(o)) + o(dt)$ and

$$\text{Var}(dA(t)) = \frac{2kA^2(t)}{A(o) + B(o)} dt + o(dt).$$

One may therefore approximate the behavior of the A process using the s.d.e.

$$dA(t) = -\frac{kA^2(t)}{N} dt + \left(\frac{2kA^2(t)}{N}\right)^{1/2} dW(t) \quad 3.A.2$$

where: $A(0) + B(0) = N$.

Following the methods of section 2, the process $X_N(t) = (A(t) - Na(t))/N^{1/2}$ is introduced, and using Ito's Lemma the s.d.e. governing $\{X_N(t), t \geq 0\}$ is found to be

$$dX_N(t) = -N^{1/2}(a'(t) + ka^2(t))dt - 2ka(t)X_N(t)dt + (2ka^2(t) + O(N^{-1/2}))^{1/2} dW(t). \quad 3.A.3$$

The coefficient of $N^{1/2}$ must be identically 0, thus $a(t)$ must satisfy

$$a'(t) = -ka^2(t), \quad a(0) = A(0)/(A(0) + B(0)) \quad \text{or} \quad a(t) = \frac{a(0)}{1 + a(0)kt}. \quad 3.A.4$$

Assuming $a(t)$ satisfies (3.A.4) and $N \rightarrow \infty$, $\{X_N(t), t \geq 0\}$ converges weakly to $\{X(t), t \geq 0\}$ which satisfies

$$dX(t) = -2ka(t)X(t) + (2ka^2(t))^{1/2} dW(t) \quad 3.A.5$$

$$X(0) = 0.$$

Once again the diffusion approximation gives a representation for $A(t)$ as $(A(0) + B(0))(a(t) + X(t))$ where $a(t)$ is given by (3.A.4). The X process is a nonstationary Ornstein-Uhlenbeck process with variance Σ_t satisfying

$$\dot{\Sigma}_t = -4ka(t)\Sigma_t + 2ka^2(t), \quad \Sigma_0 = 0. \quad 3.A.6$$

Equation (3.A.6) is easily solved to give

$$\Sigma_t = \frac{2}{3} a(o)((1+a(o)kt)^3 - 1)/(1+a_o kt)^4. \quad 3.A.7$$

It follows that marginally $A(t)$ will have approximately a normal distribution with mean $(A(o) + B(o))a(t)$ and variance $(A(o) + B(o))\Sigma_t$ where $a(t)$ and Σ_t are given in (3.A.4) and (3.A.7).

The dynamic behavior of the A process can be easily deduced from its characterization as an Ornstein-Uhlenbeck process. The simplicity of this characterization should be contrasted with the analysis presented by McQuarrie^[1]. Only the moments are presented, and they appear as complicated combinations of gamma functions.

It should be noted that the stochastic formulation given in (3.A.1) differs from that given by all other researchers with the presence of the $A(o) + B(o)$ factor in the denominator of the transition rates. This factor is important when dealing with non-linear reactions. First, this denominator makes the units of the rate reaction constants the same in the unimolecular and bimolecular cases, whereas they would be different if the transition rates were taken to be $kA^2(t)dt$. Second, it must be the case that as the reaction proceeds, B molecules are produced and this makes it difficult for the A molecules to be paired up. Using the model without the $A(o) + B(o)$ factor, the time for this reaction to be 50% complete, T_1 , is $1/2A(o)$. If the initial number of A molecules is doubled, the time for this reaction to be 75% complete, T_2 , is $3/8A(o)$. In fact the latter time should be larger than the former, since for the latter to be 75% complete, it must

first be 50% complete (at which time there are $A(o)$ A molecules left). This takes time $T_3 = 1/4A(o)$. After reaching 50% completion it takes an additional T_4 time units to reach 75% completion. Unfortunately $T_4 = T_1$ so T_2 should exceed T_1 which leads to a clear contradiction. Throughout this paper, we normalize all nonlinear transition rates by the number of molecules present. This will equalize the transition rates when both linear and nonlinear terms are included (see section 3C).

B. The Irreversible Reaction $A + B \rightarrow C$ [1],[18],[20]

The quantity $A(t)$ again represents the number of A molecules at time t . We let $A(o) + B(o) = N$ and note that $B(t) = B(o) - A(o) + A(t)$ represents the number of B molecules present at time t . The transitions for the Markov chain $\{A(t), t \geq 0\}$ are given by

<u>Transitions</u>	<u>Rate</u>
$a \rightarrow a - 1$	$ka(B(o) - A(o) + a)dt/N + o(dt)$
$a \rightarrow a$	$1 - ka(B(o) - A(o) + a)dt/N + o(dt)$
all others	$o(dt)$.

3.B.1

The approximate s.d.e. for the A process is given by

$$dA(t) = -k \frac{A(t)}{N} (B(o) - A(o) + A(t))dt + \left(k \frac{A(t)}{N} (B(o) - A(o) + A(t)) \right)^{1/2} dW(t).$$

3.B.2

The process $X_N(t) = (A(t) - Na(t))/N^{1/2}$ will also be a diffusion process, and using Ito's Lemma the s.d.e. for the $\{X_N(t), t \geq 0\}$ process is found to be

$$\begin{aligned}
dX_N(t) = & -N^{1/2}(a'(t) + ka(t)(2b - 1 + a(t)))dt \\
& - k(2b - 1 + 2a(t))X_N(t)dt \\
& + (ka(t)(2b - 1 + a(t)) + O(N^{-1/2}))^{1/2}dw(t)
\end{aligned}
\tag{3.B.3}$$

where $b = B(0)/(A(0) + B(0))$.

Letting $N \rightarrow \infty$ and assuming $a(t)$ satisfies

$$a'(t) = -ka(t)(2b - 1 + a(t)), \quad a(0) = 1 - b \tag{3.B.4}$$

then $\{X_N(t), t \geq 0\}$ converges weakly to $\{X(t), t \geq 0\}$ which satisfies

$$\begin{aligned}
dX(t) = & -k(2b - 1 + 2a(t))X(t)dt + (ka(t)(2b - 1 + a(t)))^{1/2}dw(t) \\
X(0) = & 0.
\end{aligned}
\tag{3.B.5}$$

The deterministic function $a(t)$ is easily found to be

$$a(t) = \begin{cases} \frac{(2b - 1)(1 - b)}{b \exp((2b - 1)kt) - (1 - b)}, & 0 < b < 1, b \neq 1/2 \\ 1/(2 + kt) & , \quad b = 1/2 \end{cases}
\tag{3.B.6}$$

The X process is nonstationary Ornstein-Uhlenbeck, hence the dynamic behavior can be easily deduced. The mean will be 0 for all t , and the variance, Σ_t , will satisfy

$$\begin{aligned}
\dot{\Sigma}_t = & -2k(2b - 1 + 2a(t))\Sigma_t + ka(t)(2b - 1 + a(t)) \\
\Sigma_0 = & 0.
\end{aligned}
\tag{3.B.7}$$

This equation can be easily solved using (3.B.6) to give

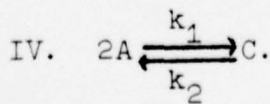
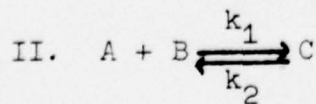
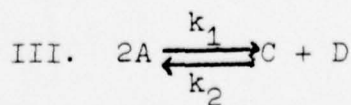
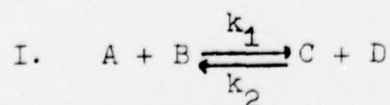
$$\Sigma_t = \begin{cases} \frac{1}{3} a(t) & b = \frac{1}{2} \\ [a(t)(\theta + a(t))]^2 \left[\frac{\theta + 2a(t)}{\theta^2 a(t)(\theta + a(t))} - \frac{\theta + 2a(o)}{\theta^2 a(o)(\theta + a(o))} \right. \\ \quad \left. - \frac{2}{3} \log \left(\frac{\theta + a(t)}{a(t)} \cdot \frac{a(o)}{\theta + a(o)} \right) \right] & 0 < b < 1 \\ & b \neq 1/2 \end{cases} \quad 3.B.8$$

$$\theta = 2b - 1$$

The stochastic evolution of A is thus given in terms of an Ornstein-Uhlenbeck process. For large values of A(o) and B(o), A(t) will be approximately normally distributed with mean (A(o) + B(o))a(t) and variance (A(o) + B(o)) Σ_t where a(t) and Σ_t are given above. The simplicity of these results should be contrasted with the complexity of the exact analysis.

C. Reversible Biomolecular Reactions [1],[13],[18],[21],[22],[23],[24]

The results given in sections 3A and 3B can be easily extended to the case where the reaction is reversible. In the case of a reversible reaction one may also calculate an equilibrium or steady state distribution. As usual A(t) represents the concentration of A molecules present at time t, and the concentration of all other types of molecules can be determined from A(t) and the initial conditions. We consider four types of reversible reactions



The method of diffusion approximations outlined in the previous sections can be routinely applied to each of the four reaction types given above. The details of the derivation are omitted as they are identical to those given in section 3A and 3B. It is, however, important to note that all nonlinear transition rates will be normalized by N , the total number of molecules present at time 0. This is especially important in cases II and IV. If, in II, we assume $A(0)$ and $B(0)$ are of a comparable order of magnitude, then the unnormalized forward transition rate will be an order of magnitude larger than the backward rate. Consequently, the reaction will, for large values of $A(0)$, be essentially equivalent to the irreversible reaction. The equilization of transition rates is accomplished by normalization.

The process $\{A(t), t \geq 0\}$ is taken to be a Markov process with appropriate transition rates. The diffusion approximation methodology treats $A(t) = Na(t) + N^{1/2}X(t)$ where N is the number of molecules of all types at time 0, $a(t)$ is a deterministic function, and $X(t)$ is a diffusion process satisfying a s.d.e. We write

$$a'(t) = r_1 a^2(t) - r_2 a(t) + r_3, \quad a(0) = a_0 \quad 3.C.1$$

$$\begin{aligned} dX(t) &= -g(t)X(t)dt + (h(t))^{1/2}dW(t) \\ X(0) &= 0, \end{aligned} \quad 3.C.2$$

and let $a_0 = A(0)/N$, $b_0 = B(0)/N$, $c_0 = C(0)/N$, and $d_0 = D(0)/N$.

We summarize below the definitions of r_1 , r_2 , r_3 , g , and h for each of the four types of reactions.

	r_1	r_2	r_3
I	$k_2 - k_1$	$k_1(b_0 - a_0) + k_2(c_0 + d_0 + 2a_0)$	$k_2(c_0 + a_0)(d_0 + a_0)$
II	$-k_1$	$k_1(b_0 - a_0) + k_2$	$k_2(c_0 + a_0)$
III	$\frac{k_2}{2} - k_1$	$k_2(c_0 + d_0 + a_0)$	$\frac{k_2}{2}(2c_0 + a_0)(2d_0 + a_0)$
IV	$-k_1$	k_2	$k_2(2c_0 + a_0)$

	$g(t)$	$h(t)$
I	$k_1(b_0 - a_0 + 2a(t))$ $+ k_2(c_0 + d_0 + 2a_0 - 2a(t))$	$k_1a(t)(b_0 - a_0 + a(t))$ $+ k_2(c_0 + a_0 - a(t))(d_0 + a_0 - a(t))$
II	$k_1(b_0 - a_0 + 2a(t)) + k_2$	$k_1a(t)(b_0 - a_0 + a(t)) + k_2(c_0 + a_0 - a(t))$
III	$2k_1a(t)$ $+ k_2(c_0 + d_0 + a_0 - a(t))$	$2k_1a^2(t)$ $+ k_2(2c_0 + a_0 - a(t))(2d_0 + a_0 - a(t))$
IV	$2k_1a(t) + k_2$	$2k_1a^2(t) + 2k_2(2c_0 + a_0 - a(t))$

The mean of $A(t)$ is obtainable as $Na(t)$ where $a(t)$ is the solution of (3.C.1). The steady state is obtained by solving (3.C.1) with $a'(t) = 0$ and $a(t) = a$. The variance of $A(t)$, Σ_t , is obtained by solving

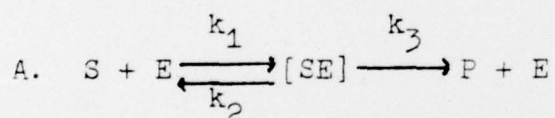
$$\dot{\Sigma}_t = -2g(t)\Sigma_t + h(t), \quad \Sigma_0 = 0. \quad 3.C.3$$

In all cases $A(t)$ will be a Gaussian process. This

characterization is especially appealing when compared with the exact answers where the moments alone require complicated combinations of hypergeometric functions.

4. Enzyme Kinetic Models

The last class of kinetic models to be addressed are those relating to enzyme kinetics. Two simple enzyme reactions are studied, one which leads to the classical Michaelis-Menten or Briggs-Haldane theory and a second which includes enzyme inhibitors. It is the purpose of this section to provide a new stochastic analysis of these reactions one which may be compared to that given by Bartholomay^[25] and Jachimowski et al^[26]. The stochastic process (now multivariate) will again be an Ornstein-Uhlenbeck process. The moments of the reaction and the n-dimensional distributions will be characterized exactly.



The simple enzyme reaction given above gives rise to the Michaelis-Menten or Briggs-Haldane theory which postulate on intermediate $[SE]$ complex. Let $S(t)$ represent the substrate concentration at time t , $E(t)$ represent the enzyme concentration at time t , and $C(t)$ and $P(t)$ represent the concentrations of enzyme-substrate complex and products at time t .

$\{(S(t), E(t)), t \geq 0\}$ is taken to form a Markov chain since $C(t) = E(0) - E(t)$ and $P(t) = S(0) - S(t) - C(t)$.

<u>Transition</u>	<u>Rate</u>
$(S, E) \rightarrow (S-1, E-1)$	$\frac{k_1 SE}{E(0)} dt + o(dt)$

TransitionRate

(continued)

4.A.1

(S+1, E+1)	$k_2(E(o) - E)dt + o(dt)$
(S, E+1)	$k_3(E(o) - E)dt + o(dt)$
(S, E)	$1 - \left(\frac{k_1 SE}{E(o)} + (k_2 + k_3)(E(o) - E) \right) dt + o(dt)$
all others	$o(dt).$

The transitions and rates given in (4.A.1) can be used to give the s.d.e. representation of the reaction. Specifically

$$\begin{aligned}
 dS(t) &= -k_1 \frac{S(t)E(t)}{E(o)} dt + k_2(E(o) - E(t))dt & 4.A.2 \\
 &\quad - \left(\frac{k_1 S(t)E(t)}{E(o)} \right)^{1/2} dW_1(t) + \left(k_2(E(o) - E(t)) \right)^{1/2} dW_2(t) \\
 dE(t) &= \frac{k_1 S(t)E(t)}{E(o)} dt + k_2(E(o) - E(t))dt + k_3(E(o) - E(t))dt \\
 &\quad - \left(\frac{k_1 S(t)E(t)}{E(o)} \right)^{1/2} dW_1(t) + \left(k_2(E(o) - E(t)) \right)^{1/2} dW_2(t) \\
 &\quad + \left(k_3(E(o) - E(t)) \right)^{1/2} dW_3(t).
 \end{aligned}$$

We next introduce $X_N(t) = (S(t) - Ns(t))/N^{1/2}$ and $Y_N(t) = (E(t) - Ne(t))/N^{1/2}$ where $N = E(o)$. The s.d.e.'s governing, $\{(X_N(t), Y_N(t)), t \geq 0\}$ can be easily deduced using the Lemma to be

$$\begin{aligned}
 dX_N(t) &= -N^{1/2}(s'(t) + k_1 s(t)e(t) - k_2(1 - e(t)))dt \\
 &\quad - k_1(s(t)Y_N(t) + e(t)X_N(t))dt - k_2Y_N(t)dt \\
 &\quad - (k_1 s(t)e(t) + o(N^{-1/2}))^{1/2} dW_1(t) & 4.A.3 \\
 &\quad + (k_2(1 - e(t)) + o(N^{-1/2}))^{1/2} dW_2(t)
 \end{aligned}$$

$$\begin{aligned}
dY_N(t) = & -N^{1/2}(e'(t) + k_1 s(t)e(t) - (k_2 + k_3)(1 - e(t)))dt \\
& - k_1(s(t)Y_N(t) + e(t)X_N(t))dt - (k_2 + k_3)Y_N(t)dt \\
& - (k_1 a(t)e(t) + O(N^{-1/2}))^{1/2} dW_1(t) \\
& + (k_2(1 - e(t)) + O(N^{-1/2}))^{1/2} dW_2(t) \\
& + (k_3(1 - e(t)) + O(N^{-1/2}))^{1/2} dW_3(t).
\end{aligned}$$

The analysis proceeds by letting $N \rightarrow \infty$ and requiring $s(t)$ and $e(t)$ to satisfy

$$\begin{aligned}
s'(t) &= -k_1 s(t)e(t) + k_2(1 - e(t)) \\
e'(t) &= -k_1 s(t)e(t) + (k_2 + k_3)(1 - e(t))
\end{aligned}
\tag{4.A.4}$$

$$e(0) = 1, \quad s(0) = S(0)/E(0).$$

Of course (4.A.4) can be augmented by the equation $c'(t) = k_1 e(t)s(t) - (k_2 + k_3)(1 - e(t))$, and this system of differential equations is now a normalized version of the Michaelis-Menten equations. The Michaelis-Menten theory can thus be thought of as a first order approximation of the enzyme reaction. The second order approximation arises from an analysis of the stochastic term.

If $s(t)$ and $e(t)$ satisfy (4.A.4), then as $N \rightarrow \infty$, $\{(X_N(t), Y_N(t)), t \geq 0\}$ converges weakly to $\{(X(t), Y(t)), t \geq 0\}$ which is a diffusion given by

$$\begin{pmatrix} dX(t) \\ dY(t) \end{pmatrix} = -\underline{A}(t) \begin{pmatrix} X(t) \\ Y(t) \end{pmatrix} dt + B(t) \begin{pmatrix} dW_1(t) \\ dW_2(t) \\ dW_3(t) \end{pmatrix}
\tag{4.A.5}$$

where

$$\begin{pmatrix} X(0) \\ Y(0) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad \text{and}$$

$$\underline{A}(t) = \begin{pmatrix} k_1 e(t) & k_1 s(t) + k_2 \\ k_1 e(t) & k_1 s(t) + k_2 + k_3 \end{pmatrix}$$

$$\underline{B}(t) = \begin{pmatrix} -(k_1 s(t)e(t))^{\frac{1}{2}} & (k_2(1-e(t)))^{\frac{1}{2}} & 0 \\ -(k_1 s(t)e(t))^{\frac{1}{2}} & (k_2(1-e(t)))^{\frac{1}{2}} & (k_3(1-e(t)))^{\frac{1}{2}} \end{pmatrix}.$$

The equation given by (4.A.5) is a nonstationary bivariate linear equation. It is easy to show that the eigenvalues of $\underline{A}(t)$ have strictly negative real parts, hence the process is Ornstein-Uhlenbeck. The marginal distribution of $(X(t), Y(t))$ will be Gaussian with mean $(0,0)$ and covariance matrix $\underline{\Sigma}_t$ given to be the unique nonnegative definite solution of the matrix Riccati equation

$$\dot{\underline{\Sigma}}_t = -\underline{A}(t)\underline{\Sigma}_t - \underline{\Sigma}_t\underline{A}^T(t) + \underline{B}_t\underline{B}_t^T, \quad \underline{\Sigma}_0 = \underline{0} \quad 4.A.6$$

The above equation can be rewritten as follows. Let $\sigma_1(t) = \text{Var}(X(t))$, $\sigma_2(t) = \text{Var}(Y(t))$, and $\sigma_{12}(t) = \text{Cov}(X(t), Y(t))$. Then calling $\underline{V}(t) = (\sigma_1(t), \sigma_{12}(t), \sigma_2(t))^T$, we find

$$\dot{\underline{V}}(t) = -\underline{D}_t \underline{V}(t) + \underline{F}_t, \quad \underline{V}(0) = \underline{0} \quad 4.A.7$$

where

$$\underline{D}_t = \begin{pmatrix} 2k_1 e(t) & 2(k_1 s(t) + k_2) & 0 \\ k_1 e(t) & k_1(e(t) + s(t)) + k_2 + k_3 & k_1 s(t) + k_2 \\ 0 & 2k_1 e(t) & 2(k_1 s(t) + k_2 + k_3) \end{pmatrix}$$

and

$$\underline{F}_t = \begin{pmatrix} k_1 s(t)e(t) + k_2(1-e(t)) \\ k_1 s(t)e(t) + k_2(1-e(t)) \\ k_1 s(t)e(t) + (k_2 + k_3)(1-e(t)) \end{pmatrix}.$$

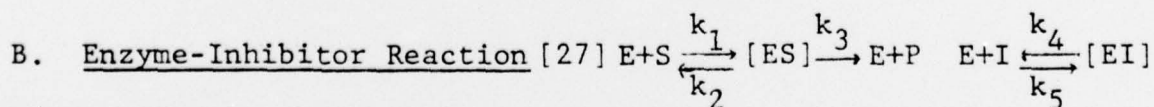
Equation (4.A.7) may be solved formally to give

$$\tilde{V}(t) = \exp(- \int_0^t D_u du) \int_0^t \exp(F_u \int_0^u D_s ds) du \quad 4.A.8$$

which can be computed using standard numerical integration techniques.

The diffusion approximation methodology thus provides a complete description of the enzyme reaction in terms of a bivariate Ornstein-Uhlenbeck process. It appears that no closed form expressions are available, so $s(t)$, $e(t)$, and $\tilde{V}(t)$ must be evaluated using numerical integration. It is possible to obtain approximations by adopting the Briggs and Haldane steady state assumption, $c'(t) = 0$. This simplification allows for the derivation of certain closed form expressions.

It should be noted that the stochastic description of this enzyme reaction is valuable not only for the insight it provides into the process, but also it is necessary for the development of a proper statistical analysis of reaction data, say to estimate the Michaelis-Menten constant $(k_2 + k_3)/k_1$. The analysis presented can be easily generalized to far more complicated reactions including reactions with several intermediate enzyme-substrate complexes, several substrates, and reversible reactions.



We conclude with a simple enzyme-inhibitor reaction. The analysis of this system parallels that given in section 4A, except that a higher dimensional process must be used. Again the diffusion

approximation analysis leads to a characterization in terms of a multivariate Ornstein-Uhlenbeck process.

Define $S(t)$, $E(t)$, and $I(t)$ to be the concentrations of substrate, enzyme, and inhibitor at time t . If $C_1(t)$ and $C_2(t)$ represent the concentrations of $[ES]$ and $[EI]$, then both can be calculated from the initial conditions by $C_2(t) = I(0) - I(t)$ and $C_1(t) = E(0) - E(t) - C_2(t)$. $\{(S(t), E(t), I(t)), t \geq 0\}$ can be treated as a Markov Chain with transitions and rates given by

<u>Transition</u>	<u>Rate</u>	
$(S, E, I) \rightarrow (S-1, E-1, I)$	$\frac{k_1 S E dt}{E(0)} + o(dt)$	
$(S+1, E+1, I)$	$k_2 (E(0) - E - C_2) dt + o(dt)$	4.B.1
$(S, E+1, I)$	$k_3 (E(0) - E - C_2) dt + o(dt)$	
$(S, E-1, I-1)$	$\frac{k_4 EI}{E(0)} dt + o(dt)$	
$(S, E+1, I+1)$	$k_5 (I(0) - I) dt + o(dt)$	

with $C_2 = I(0) - I$.

We omit writing the approximate s.d.e.'s for $(S(t), E(t), I(t))$. Let $X_N(t) = (S(t) - Ns(t))/N^{\frac{1}{2}}$, $Y_N(t) = (E(t) - Ne(t))/N^{\frac{1}{2}}$, and $Z_N(t) = (I(t) - Ni(t))/N^{\frac{1}{2}}$ where $N = E(0)$. An asymptotic analysis indicates that the deterministic functions $s(t)$, $e(t)$, and $i(t)$ must satisfy

$$\begin{aligned}
 s'(t) &= -k_1 s(t)e(t) + k_2(1-e(t) - i_0 + i(t)) \\
 e'(t) &= -k_1 s(t)e(t) + (k_2 + k_3)(1-e(t) - i_0 + i(t)) \\
 i'(t) &= -k_4 e(t)i(t) + k_5 (i_0 - i(t))
 \end{aligned}
 \tag{4.B.2}$$

where $e(0) = 1$, $I(0)/E(0) = i_0$, and $s_0 = S(0)/E(0)$.

The stochastic noise process $\{(X_N(t), Y_N(t), Z_N(t)), t \geq 0\}$ will converge weakly to $\{\underline{U}(t), t \geq 0\}$ where $\underline{U}(t) = (X(t), Y(t), Z(t))^T$ and will satisfy

$$d \underline{U}(t) = -\underline{A}_t \underline{U}(t) dt + \underline{B}_t d \underline{W}(t), \quad \underline{U}(0) = \underline{0} \quad 4.B.3$$

with $\underline{W}(t) = (W_1(t), W_2(t), W_3(t), W_4(t), W_5(t))^T$,

$$\underline{A}_t = \begin{pmatrix} k_1 e(t) & k_1 s(t) + k_2 & k_2 \\ k_1 e(t) & k_1 s(t) + k_2 + k_3 + k_4 i(t) & -(k_2 + k_3 - k_5) \\ 0 & k_4 i(t) & k_4 e(t) + k_5 \end{pmatrix},$$

$$\underline{B}_t = \begin{pmatrix} g_1(t) & g_2(t) & 0 & 0 & 0 \\ g_1(t) & g_2(t) & g_3(t) - g_4(t) & g_5(t) & \\ 0 & 0 & 0 & -g_4(t) & g_5(t) \end{pmatrix}$$

with $g_1(t) = (k_1 s(t) e(t))^{\frac{1}{2}}$, $g_2(t) = (k_2 (1 - e(t) - i_0 + i(t)))^{\frac{1}{2}}$, $g_3(t) = (k_3 (1 - e(t) - i_0 + i(t)))^{\frac{1}{2}}$, $g_4(t) = (k_4 e(t) i(t))^{\frac{1}{2}}$, and $g_5(t) = (k_5 e(t) i(t))^{\frac{1}{2}}$.

The eigenvalues of \underline{A}_t have strictly negative real parts, hence the $\underline{U}(t)$ process is Ornstein-Uhlenbeck with mean $\underline{0}$. It follows that $\underline{U}(t)$ will have a Gaussian distribution. The covariance matrix at time t , $\underline{\Sigma}_t$, is given as the unique nonnegative definite solution of (4.A.6). The diffusion approximation then gives $(S(t), E(t), I(t))$ to be approximately $E(0)(s(t), e(t), i(t)) + E(0)^{\frac{1}{2}} (X(t), Y(t), Z(t))$ where $(s(t), e(t), i(t))$ is given by (4.B.2), and the stochastic part is specified by (4.B.3). It appears impossible that $s(t), e(t), i(t)$, and $\underline{\Sigma}_t$ can be found in closed form; however, one can easily use standard numerical procedures to compute them. It

is possible that further approximations similar to the Briggs-Haldane steady state assumption can be invoked to obtain some closed form information. The joint distribution of the process over time will be given by a trivariate Gaussian distribution.

5. Acknowledgement

The author wishes to thank Professor Donald P. Gaver of the U. S. Naval Postgraduate School for introducing him to the problem of the stochastic characterization of enzyme kinetic models and for his help throughout the course of the research.

REFERENCES

- (1) McQuarrie, D. A. (1967) Stochastic Approach to Chemical Kinetics, Methuen & Co. Ltd., London.
- (2) McQuarrie, D. A. (1969) Stochastic theory of chemical rate processes. Advances in Chemical Physics XV (ed. Shuler, K. E.). Wiley-Interscience, New York, 149-184.
- (3) Singer, K. (1953) Applications of the theory of stochastic processes to the study of irreproducible chemical reactions and nucleation processes. J.R. Statist. Soc. B 15, 92-106.
- (4) Bharucha-Reid, A.T. (1960) Elements of the Theory of Markov Processes and Their Applications. McGraw-Hill, New York.
- (5) Gaver, D. P. and Lehoczky, J. P. (1976) Gaussian approximations to service problems: a communication system example. J. Appl. Prob. 13, 768-780.
- (6) Gaver, D. P. and Lehoczky, J. P. (1977) A diffusion approximation for a repairman problem with two types of repair. To appear Management Science.
- (7) Lehoczky, J. P. and Gaver, D. P. (1977) A diffusion approximation analysis of a general n-compartment system. To appear Math. Biosciences.
- (8) McNeil, D. R. and Schach, S. (1973) Central limit analogues for Markov population processes. J. R. Statist. Soc. B 35, 1-23.
- (9) Arnold, L. (1973) Stochastic Differential Equations: Theory and Applications. John Wiley & Sons, New York.
- (10) Gikhman I. I. and Skorokhod, A. V. (1971) Stochastic Differential Equations. Springer-Verlag, New York.
- (11) Bartholomay, A. F. (1958) Stochastic models for chemical reactions, I. Theory of the unimolecular reaction process. Bull. Math. Biophys. 20, 175-190.

- (12) Bartholomay, A. F. (1959) Stochastic models for chemical reactions, II. The unimolecular rate constant. Bull. Math. Biophys. 21, 363-373.
- (13) McQuarrie, D. A. (1963) Kinetics of small systems, I. J. Chem. Phys. 38, 433-436.
- (14) Solc, M. and Horsak, I. (1972) Simulation of first-order chemical reaction as a stochastic process on a digital computer. Coll. Czech. Chem. Comm. 37, 2994-3000..
- (15) Kurtz, T. G. (1971) Limit theorems for sequences of jump Markov processes approximating ordinary differential processes. J. Appl. Prob. 8, 344-356.
- (16) Barbour, A. (1974) On a functional central limit theorem for Markov population processes. Adv. Appl. Prob. 6, 21-39.
- (17) Horsak, I. and Solc, M. (1973) Simulation study of a stochastic model of reversible first-order reaction equilibrium. Coll. Czech. Chem. Comm. 38, 2200-2203.
- (18) McQuarrie, D. A., Jachinowski, C. J., and Russell, M. E. (1964) Kinetics of small systems, II. J. Chem. Phys. 40, 2914-2921.
- (19) Solc, M. (1972) Alternate stochastic formulation of first order reaction kinetics. Coll. Czech. Chem. Comm. 37, 1055-1060.
- (20) Ishida, K. (1964) Stochastic model for bimolecular reaction. J. Chem. Soc. 41, 2472-2478.
- (21) Darvey, I. G., Ninham, B. W., and Staff, P. J. (1966) Stochastic models for second order chemical reaction kinetics. The equilibrium state. J. Chem. Phys. 45, 2145-2155.
- (22) Darvey, I. G. and Ninham, B. W. (1967) Stochastic models for second-order chemical reaction kinetics. Time course of reactions. J. Chem. Phys. 46, 1626-1645.

- (23) Solc, M. (1974) Stochastic model of reversible bimolecular equilibrium. Coll. Czech. Chem. Comm. 39, 197-205.
- (24) Staff, P. J. (1967) Approximation method for equilibrium distributions in second-order chemical reaction kinetics. J. Chem. Phys. 46, 2209-2212.
- (25) Bartholomay, A. F. (1962) A stochastic approach to statistical kinetics with application to enzyme kinetics. Biochemistry 1, 223-230.
- (26) Jachimowski, C. J., McQuarrie, D. A., and Russell, M. E. (1964) A stochastic approach to enzyme-substrate reactions. Biochemistry 3, 1732-1736.
- (27) Webb, J. L. (1963) Enzyme and Metabolic Inhibitors Vol. 1, Academic Press, New York.

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM	
1. REPORT NUMBER AFOSR-TR-77-1212	2. GOVT ACCESSION NO. 91 Interim Rept.	3. RECIPIENT'S CATALOG NUMBER	
4. TITLE (and Subtitle) A NEW STOCHASTIC ANALYSIS OF CHEMICAL KINETICS	5. TYPE OF REPORT & PERIOD COVERED Interim	6. PERFORMING ORG. REPORT NUMBER Tech Report No 129	
7. AUTHOR(s) John P. Lehoczky	8. CONTRACT OR GRANT NUMBER(s) AFOSR-74-2642	9. PERFORMING ORGANIZATION NAME AND ADDRESS Carnegie-Mellon University Department of Statistics Pittsburgh, PA 15213	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 61102F 2304/A5
11. CONTROLLING OFFICE NAME AND ADDRESS Air Force Office of Scientific Research/NM Bolling AFB, Washington, DC 20332	12. REPORT DATE June 1977	13. NUMBER OF PAGES 28	14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office) 1233p.
15. SECURITY CLASS. (of this report) UNCLASSIFIED		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE	
16. DISTRIBUTION STATEMENT (of this Report) Approved for public release; distribution unlimited.			
17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)			
18. SUPPLEMENTARY NOTES			
19. KEY WORDS (Continue on reverse side if necessary and identify by block number) chemical reaction, stochastic model, chemical kinetics, Ornstein-Uhlenbeck process, diffusion approximation, enzyme kinetic, Michaelis-Menten theory			
20. ABSTRACT (Continue on reverse side if necessary and identify by block number) A new approach to the formulation and analysis of stochastic models of chemical reactions is presented. Unimolecular, bimolecular, and enzyme kinetic reactions are considered in the irreversible and reversible cases. The methodology is based on diffusion approximations and represents the time evolution of the reaction as the sum of a deterministic function and an Ornstein-Uhlenbeck process. As a result the marginal distributions are approximately Gaussian with relatively simple mean and covariance parameters, and the dynamic behavior is completely characterized. The stochastic approach which uses stochastic			

20. Abstract

→ differential equations is a natural generalization of the deterministic approach which uses ordinary differential equations. ↑

UNCLASSIFIED